MIT LIBRARIES DUPL 1
3 9080 00578985 1
HB31
.M415
no.535

MASSACHUSETTS
INSTITUTE of TECHNOLOGY
LIBRARIES

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/permanenttransit00quah

HB31
.M415
no.535

working paper
department
of economics

DEC 21 19

PERMANENT AND TRANSITORY MOVEMENTS IN LABOR INCOME: AN EXPLANATION FOR "EXCESS SMOOTHNESS" IN CONSUMPTION

Danny Quah

No. 535 September 1989

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

PERMANENT AND TRANSITORY MOVEMENTS IN LABOR INCOME: AN EXPLANATION FOR "EXCESS SMOOTHNESS" IN CONSUMPTION

Danny Quah

No. 535 September 1989

DEC 21 1989
RECEIVED

Permanent and Transitory Movements in Labor Income: An Explanation for "Excess Smoothness" in Consumption.

Danny Quah

Economics Department, MIT.
September 1989.

Abstract

Many have argued that if labor income is difference stationary, the permanent income hypothesis predicts consumption should be relatively volatile. In US aggregate data, labor income is well-characterized as having a unit root; however consumption turns out to be relatively smooth. This anomaly is known as Deaton's paradox. We resolve Deaton's paradox by providing decompositions of labor income into permanent and transitory components. These preserve the univariate dynamic properties of labor income. However, when agents distinguish permanent and transitory movements in their labor income—as the rational expectations hypothesis asserts they should—the permanent income hypothesis correctly predicts the observed smoothness in consumption.

1. Introduction

Milton Friedman's permanent income theory of consumption is one of the outstanding successes of dynamic economic reasoning. This theory asserts that consumption occurs out of permanent income, not current income. Permanent income is related to but distinct from current observed income. Under the intuition that permanent income—because it is "permanent"—should be smoother than current income, the theory has long been understood to predict that consumption should be smooth relative to observed income fluctuations. This relative smoothness of consumption is a firmly established empirical regularity in aggregate time series data.

Deaton [1987] observed, however, that the permanent income hypothesis fails to generate this smoothness if labor income is an *integrated process*, i.e., if labor income has a unit root. According to Deaton, a unit root characterization for labor income, given the data, implies that observed consumption is insufficiently sensitive to innovations in current income. Deaton concluded that if labor income is well-characterized as being integrated, then "the representative agent version of the permanent-income hypothesis can be rejected because it fails to predict the fact that consumption is smooth, the very fact that it was invented to explain in the first place" (Deaton [1987, p. 122]). This anomaly in the joint behavior of consumption and income has come to be known as *Deaton's paradox*.

Deaton's analysis, therefore, appears to argue strongly for the need to establish whether labor income truly is an integrated process. His work has suggested that the predictions of an important economic theory—Friedman's permanent income hypothesis—are intimately related to measures of long run persistence, such as considered by Campbell

and Mankiw [1987] and Cochrane [1988].[1]

This paper offers a simple and intuitive explanation for this smoothness in consumption. There are different kinds of disturbances that impinge on the labor income stream. Some disturbances have permanent effects on labor income; other disturbances have only transitory impact. I show below that the permanent income hypothesis under rational expectations—not surprisingly—implies different kinds of disturbances have different effects on consumption. Disturbances that do not have permanent effects on labor income will not have large effects on Friedman's permanent income. These disturbances will therefore have only relatively small impact on consumption. On the other hand, disturbances that do have a permanent impact on labor income will have relatively large effects on consumption.[2]

Therefore, according to the analysis here, the permanent income hypothesis prediction for the smoothness properties of consumption depends on the relative importance of permanent and transitory components in labor income. The univariate dynamics of labor income—whether or not labor income is integrated, or how "persistent" labor income is, or the precise form of the univariate dynamics—turn out to be not especially informative for the predictions of the theory.

It remains controversial whether macroeconomic time series are better characterized as being integrated or as being stationary about a deterministic trend. This paper does not attempt to shed light on that issue. Instead it argues only that at least within the context of the permanent income hypothesis, a unit root characterization for labor income may not have implications that are as dramatic as has previously been suggested. Further, and

again at least within the context of the permanent income hypothesis, the widely-noted measures of persistence in Campbell and Mankiw [1987] and Cochrane [1988] may simply be beside the point.[3]

The remainder of this paper is organized as follows. Section 2 briefly reviews other explanations of "excess smoothness" that have been offered, and makes explicit the difference between those and the reasoning in this paper. Section 3 sets out the standard permanent income model, and makes rigorous the intuition above. Section 4 provides decompositions of labor income into permanent and transitory components that reconcile the following: (1) the observed smoothness in aggregate consumption, (2) the estimated univariate dynamics of labor income maintaining a unit root characterization, and (3) the permanent income hypothesis.[4] Our explanation for apparent "excess smoothness" in consumption turns on the plausible assumption that economic agents forecast future labor income using strictly more information than does the econometrician. In many rational expectations models, the econometrician can take into account this superior information of private agents by using the endogenous variables of the model in the modeller's forecasting equations. Section 5 shows the permanent income model of consumption to be a counter-example to the validity of this methodology: A researcher studying the observed behavior of the model variables would conclude that consumption is unresponsive to "news" in labor income, even if the permanent income hypothesis were to be true.[5] Section 6 concludes the paper.

2. Related Literature

Deaton's paradox has generated an extensive literature. For reasons of space, we will only discuss a small fraction of the relevant work: Christiano [1987] and Diebold and Rudebusch [1989] provide more extended discussion on the literature.

Christiano [1987] observes that movements in labor income may be related to interest rate fluctuations. To the extent that savings are sensitive to interest rate movements, equilibrium changes in consumption will be dampened by appropriate comovements in income and the interest rate. Thus, conditional on a given pattern of labor income dynamics, an equilibrium theory might, in principle, predict consumption to be less volatile than implied by a model with a constant interest rate. Christiano therefore studies a simple general equilibrium real business cycle model that allows the interest rate to vary over time. By appropriately setting parameter values, he is able to match the observed volatility of changes in consumption. Christiano points out, however, that in doing so, the model is unable to replicate the actual dynamic behavior of income in the US economy.

Caballero [1988a] modifies the preferences of the infinitely-lived representative consumer to allow "taste shocks," and to incorporate an explicit "precautionary savings" motive. Clarida [1988] and Gali [1989] consider the aggregation problem in infinitely-lived model economies that are populated by finite-lived consumers. These modifications partially succeed in reconciling the predictions of the theory with the data. They all suggest that even in the presence of a unit root in labor income, equilibrium theory predicts that consumption may be relatively smooth. Note, however, none of these proposals quite confront the challenge that Deaton posed.[6]

This same comment applies to that class of explanations that suggest labor income may in fact not be difference stationary, or that even if it were difference stationary, the usual estimates of long-run persistence may simply be "too large." Diebold and Rudebusch [1989] suggested a fractional integration model for labor income; Watson [1986] used an unobserved components model. In US aggregate data, these alternative parametrizations of the Wold moving average representation imply point estimates for long run persistence smaller than those in Deaton [1987] or West [1988]. Again, these suggestions partially succeed in reconciling the optimizing theory with the data. As does Cochrane [1988], these papers properly warn that estimates of long run persistence may be quite sensitive to specification. However, according to the analysis developed below, the estimates of long run persistence are simply not especially relevant.

The conclusions of this paper rely on the researcher having strictly less information than agents. In many rational expectations applications, this is not important as the model variables will reveal all relevant information. This insight underlies the many Euler equation-type tests of market efficiency and equilibrium models. In the current setting however, when there is more than one disturbance affecting labor income, the permanent income hypothesis also predicts that the model variables, consumption and income, can *not* appropriately reveal the true effects on consumption of "news" in labor income. In fact, an example below shows that an econometrician studying the joint dynamics of consumption and income will conclude that consumption seems not to respond to certain "news" in labor income—even when the permanent income hypothesis is true. Thus, the econometrician will conclude that consumption appears to be "excessively smooth," even though in truth,

it isn't. Note however that this does not explain the rejections of the permanent income hypothesis in Campbell and Deaton [1989] and West [1988], as those researchers employed information on asset holdings, in addition to consumption and income. Following the reasoning in Campbell and Deaton, that rejection must therefore arise from violation of the usual cross equation restrictions, and *not* from "excess smoothness" per se.

Flavin [1988] has criticized the work by Campbell and Deaton [1989] and West [1988] from a different perspective. Flavin's model departs from the permanent income theory; under the hypothesis in that work, consumption and savings will not contain all relevant information. By contrast, we argue here that the main force of Flavin's conclusion holds *even under the permanent income hypothesis.*

3. The Model

Hansen [1987] and Sargent [1989] have provided dynamic general equilibrium interpretations of the permanent income hypothesis (hereafter PIH). The specification follows that in Hall [1978] and Flavin [1981], and comprises the three equations:

$$C(t) = rW(t), \tag{3.1}$$

$$W(t) = K(t) + \left[(1+r)^{-1}\sum_{j=0}^{\infty}(1+r)^{-j}E_tY(t+j)\right], \tag{3.2}$$

$$K(t+1) = (1+r)K(t) + Y(t) - C(t). \tag{3.3}$$

Equation (3.1) states consumption in each period equals *permanent income*—this is simply the flow of rental income from total wealth W, accruing at the time-invariant equilibrium risk-free interest rate r. *Total wealth* is the sum (3.2) of physical capital K and human

wealth. *Human wealth*, in turn, is the expected present discounted value of the stream of labor income Y. As usual in this literature, we assume the labor income stream is exogenous with respect to the agent's consumption decision. However, total income—the sum of labor and capital income—obviously depends on the agent's past consumption decisions. In summary, the agent consumes the resource stream that flows from renting out, in perfect markets, all of her physical and human capital. Equation (3.3) simply defines capital stock transition: capital doesn't depreciate, and accumulates through agents' savings.

These equations can be combined to obtain:

$$\Delta C(t) = C(t) - C(t-1) = \frac{r}{1+r} \cdot \sum_{j=0}^{\infty} (1+r)^{-j} \left(E_t Y(t+j) - E_{t-1} Y(t+j) \right).$$

Defining β to equal $(1+r)^{-1}$, this is:

$$\Delta C(t) = (1-\beta) \sum_{j=0}^{\infty} \beta^j \left(E_t Y(t+j) - E_{t-1} Y(t+j) \right). \tag{3.4}$$

As numerous authors have emphasized, the right hand side of (3.4) is the annuity value of revisions in the expected labor income stream—these revisions are due to new information arriving in period t. The model predicts the larger is the impact of "news" on human wealth, the larger should be the change in consumption.

Equation (3.4) yields two potentially refutable predictions. First, given a particular data generating process for labor income, the magnitude of consumption's response to news can be calculated from (3.4). Second, information available prior to an arbitrary time period t should neither affect nor help to predict the change in consumption at t, i.e., consumption should be a martingale with respect to agents' information—this is simply Hall's [1978] famous characterization of consumption under the PIH. Contradiction of these implications is referred to as "excess smoothness" and "excess sensitivity" respectively.

In this model, Hall's martingale characterization is clearly independent of the exact process that generates labor income. However, the appropriate statistical theory for inference should, of course, depend on the properties of the instruments used to examine the martingale restriction. But this will always be true in any econometric procedure, and is not particularly special to the PIH.

The "smoothness" predictions, however, depend critically on the model generating labor income. That model is what defines *news* which, in turn, affects consumption through (3.4). To see this explicitly, we briefly summarize Deaton's [1987] excess smoothness argument.

First, suppose labor income Y is a trend-stationary process. We can, without loss, take the trend to be identically zero, since we are here interested only in the second moment properties of consumption and income. If Y has finite time-invariant second moments, it is guaranteed to have a unique Wold representation:

$$Y(t) = \sum_{k=0}^{\infty} b(k)\eta(t-k) = B(L)\eta(t),$$

where $b(0) = 1$, the function $B(z) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} b(k)z^k \neq 0$ for all z on the closed unit disk, L denotes the lag operator, and η is serially uncorrelated. Suppose the representative agent uses only current and lagged labor income observations to forecast future labor income. A result due to Hansen and Sargent [1980, Appendix A] then implies a simple formula for human wealth:[7]

$$\sum_{j=0}^{\infty} \beta^j E_t Y(t+j) = \left[\frac{LB(L) - \beta B(\beta)}{L - \beta}\right] \eta(t).$$

Further, by a property of iterated expectations:

$$\sum_{j=0}^{\infty}\beta^j E_{t-1}Y(t+j) = E_{t-1}\left(\sum_{j=0}^{\infty}\beta^j E_t Y(t+j)\right) = E_{t-1}\left(\left[\frac{LB(L)-\beta B(\beta)}{L-\beta}\right]\eta(t)\right)$$
$$= \left[\frac{LB(L)-\beta B(\beta)}{L-\beta}L^{-1}\right]_+ L\eta(t),$$

where $[\quad]_+$ denotes the annihilation operator.[8] Using these in (3.4), the resulting change in consumption is:

$$\Delta C(t) = (1-\beta)\left(\frac{LB(L)-\beta B(\beta)}{L-\beta} - \left[\frac{LB(L)-\beta B(\beta)}{L-\beta}L^{-1}\right]_+ L\right)\eta(t).$$

This simplifies to:

$$\Delta C(t) = (1-\beta)B(\beta)\eta(t). \tag{3.5}$$

Thus, in any given period, the change in consumption depends on (1) the interest rate, through β, (2) the dynamics B of labor income, and (3) the innovation η in labor income. Given β and B, the change in consumption is proportional to news in labor income. Since β is close to 1 for small values of the interest rate, other things equal, changes in consumption should be relatively "small."

Next, suppose labor income is difference stationary. As in our treatment of the trend stationary case, we ignore possible drift in labor income, since that cannot affect the second moment properties of consumption and income. Denote changes in Y by ΔY. Assuming the process ΔY has finite time-invariant second moments, it necessarily has a unique Wold representation:

$$\Delta Y(t) = \sum_{k=0}^{\infty} a(k)\epsilon(t-k) = A(L)\epsilon(t),$$

where $a(0) = 1$, the function $A(z) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} a(k)z^k \neq 0$ for $|z| \leq 1$, and ϵ is serially uncorrelated. If agents use only current and lagged values of labor income to forecast

future labor income—as Deaton [1987] assumed—we can again use the Hansen-Sargent analysis to obtain:[9]

$$\sum_{j=0}^{\infty} \beta^j E_t Y(t+j) = \left[\frac{LA(L) - \left(\frac{\beta}{1-\beta}\right) A(\beta)(1-L)}{(L-\beta)(1-L)} \right] \epsilon(t).$$

As before, a property of iterated expectations implies that:

$$\sum_{j=0}^{\infty} \beta^j E_{t-1} Y(t+j) = \left[\frac{LA(L) - \left(\frac{\beta}{1-\beta}\right) A(\beta)(1-L)}{(L-\beta)(1-L)} L^{-1} \right]_+ L\epsilon(t).$$

By the same reasoning as above, changes in consumption follow:

$$\Delta C(t) = (1-\beta) \cdot (1-\beta)^{-1} A(\beta)\epsilon(t) = A(\beta)\epsilon(t). \tag{3.6}$$

Equation (3.6) is the analogue to (3.5) when labor income is difference stationary rather than trend stationary. Comparing these two equations, notice (3.6) does not contain the term $(1-\beta)$; this is why, other things equal, the PIH under difference stationarity implies a relatively more volatile consumption series.

When A and $\mathrm{Var}(\epsilon)$ are estimated on US aggregate time series data, the implied variance of the right hand side of (3.6) is significantly larger than the sample variance of changes in consumption.[10] From this evidence, Deaton [1987] concluded that aggregate consumption is excessively smooth if labor income is characterized as an integrated process.

In US aggregate data, labor income appears to be well-described as being integrated.[11] However, there is certainly no compelling evidence that agents in the economy estimate human wealth using only their labor income history. For instance, suppose there are two kinds of structural disturbances to labor income. One class of disturbances has a permanent impact on the level of labor income; the other disturbances have only transitory impact. For

simplicity, we can suppose that there are only two structural disturbances in the economy—one in each class. Allowing a more general specification does not alter the conclusions of interest here, although typically the disturbances will not aggregate naturally into the one permanent and one transitory component (see the Technical Appendix in Blanchard and Quah [1989]).

Under rational expectations, agents estimate their human wealth using all available information on the different kinds of disturbances. In particular, if it will improve their forecasts of future labor income, they will use the information that there is a permanent and a transitory component in labor income. The innovations in the different structural disturbances may be correlated; however, one can—in the natural way and without loss of generality—construct an orthogonal decomposition to use in forecasting future labor income.[12]

Therefore, suppose we can write:

$$Y(t) = Y_1(t) + Y_0(t),$$

where Y_1 is difference stationary, and Y_0 is covariance stationary. Assuming ΔY_1 and Y_0 have finite time-invariant second moments, we can write their Wold decompositions as:

$$\Delta Y_1(t) = \sum_{k=0}^{\infty} a_1(k)\epsilon_1(t-k) = A_1(L)\epsilon_1(t),$$

and

$$Y_0(t) = \sum_{k=0}^{\infty} a_0(k)\epsilon_0(t-k) = A_0(L)\epsilon_0(t),$$

where the innovations ϵ_1 and ϵ_0 are uncorrelated at all leads and lags.[13] For brevity, we will refer to Y_1 and Y_0 as the *permanent* and *transitory components* in labor income,

respectively. The permanent component in labor income should not be confused with *permanent income*, which is precisely defined from the equations of the PIH.

By exactly the same reasoning as in the cases previously considered, equilibrium consumption follows:

$$\Delta C(t) = A_1(\beta)\epsilon_1(t) + (1-\beta)A_0(\beta)\epsilon_0(t). \tag{3.7}$$

Equation (3.7) shows that consumption response depends, in general, on the *kind* of news that dominates in any given period. For β close to unity, news that has only transitory effects will have only relatively small impact on consumption; the opposite is true of news that turn out to have permanent effects. Thus, the theory predicts that consumption volatility depends on the relative importance of permanent and transitory components in labor income.

To see explicitly the implications for consumption volatility, we need the Wold lag distributions A_1 and A_0, as well as the innovation variances $\mathrm{Var}(\epsilon_1)$ and $\mathrm{Var}(\epsilon_0)$. While the components Y_1 and Y_0 are guaranteed to sum to the observed labor income process Y, their innovations ϵ_1 and ϵ_0 bear no simple relation to the innovation ϵ in Y. Nor is there a simple relation between the lag distributions A_1, A_0 and A: in particular, it is not true that $A_1(z) + (1-z)A_0(z)$ equals $A(z)$.

How then is such a decomposition into permanent and transitory components consistent with the time series observations on aggregate labor income? It is clearly necessary that the spectral densities of ΔY_1 and ΔY_0 sum pointwise to equal the spectral density of ΔY. Thus, for all ω in $(-\pi, \pi]$, we have:

$$\mathrm{Var}(\epsilon)|A(e^{-i\omega})|^2 = \mathrm{Var}(\epsilon_1)|A_1(e^{-i\omega})|^2 + \mathrm{Var}(\epsilon_0)|1-e^{-i\omega}|^2|A_0(e^{-i\omega})|^2.$$

Under weak regularity conditions, this relation across the spectral densities is not only necessary but also sufficient to characterize the orthogonal decomposition (see Quah [1989]). In the subsequent discussion, we can therefore focus only on this pointwise equality in the spectral densities. Making the natural definitions, we can write this as:

$$S(\omega) = S_1(\omega) + |1 - e^{-i\omega}|^2 S_0(\omega). \tag{3.8}$$

This equation has two important features that we will use repeatedly below. First, since the second term on the right hand side is nonnegative, S_1 must be everywhere bounded from above by S. Next, S_1 must equal S at $\omega = 0$, since the second term on the right hand side vanishes there. Thus, $S_1(\omega) \leq S(\omega)$ for all ω, with strict equality at $\omega = 0$. In words, the spectral density of changes in observed labor income forms an outer envelope for that in its permanent component, where that outer envelope is binding at frequency zero. Put another way, the spectral densities of changes in the permanent and transitory components is a *cleaving* of that in observed labor income.[14] Figure 1 illustrates such a cleaving of a spectral density.

By the equality at frequency zero of S and S_1, agents' forecasts of the long run effects of a disturbance are always the same—regardless of whether agents view the disturbance as one in the permanent component. or as one in observed labor income itself.[15] The relative importance of the permanent and transitory components is altered as we vary the cleaving of the spectral density. Across all cleavings, however, the measure of long run persistence always remains the same. Equation (3.7) suggests then that as long as β is strictly less than one, the volatility of consumption—which appears to vary with the relative importance of the permanent and transitory components—is not determined in any essential way by the

magnitude of long run persistence. In the next section, we will use the spectral density characterization to show this rigorously.

4. Explaining "Excess Smoothness"

A first-order autoregressive model for the first differences of US aggregate labor income yields point estimates of 0.44 for the autoregressive coefficient and 636.1 for the innovation variance. If the risk-free interest rate r is taken to be 1 percent per quarter, then equation (3.6) implies the variance of consumption changes should be 1997; the actual observed sample variance is only 246.[16] Tables 1 and 2 in West [1988] display different *ARIMA* parametrizations for labor income that all show the same conclusion. Consumption appears to be "excessively smooth," given the maintained assumption that labor income is an integrated process.

Now suppose agents use information on permanent and transitory movements in labor income to estimate their human wealth. Recall consumption should then behave as:

$$\Delta C(t) = A_1(\beta)\epsilon_1(t) + (1-\beta)A_0(\beta)\epsilon_0(t),$$

which implies the variance of consumption changes is:

$$\mathrm{Var}(\Delta C) = A_1(\beta)^2 \cdot \mathrm{Var}(\epsilon_1) + (1-\beta)^2 A_0(\beta)^2 \cdot \mathrm{Var}(\epsilon_0).$$

Clearly, the univariate Wold characterization of labor income does not directly restrict the smoothness properties of consumption.

A simple example will build intuition for the calculations to follow, although the example isn't completely successful in explaining "excess smoothness." Suppose the permanent

component Y_1 in labor income is described by $A_1(z) = (1-\gamma z)^{-1}$, with $|\gamma| < 1$, i.e., the permanent component is a stationary first order autoregressive process in first differences. Assume the correct model for observed labor income is the first order autoregression in first differences above. For the "outer envelope" condition described above to hold, we must have (1) $\gamma > 0.44$, and (2) $\mathrm{Var}(\epsilon_1) = (1-\gamma)^2 \times (1-0.44)^{-2} \times 636.1$. Condition (1) guarantees the spectral density S of ΔY dominates that of ΔY_1; condition (2) restricts these spectral densities to be equal at frequency zero. It follows, then, that $\Delta Y - \Delta Y_1$ is the first difference of a process that is covariance stationary.[17]

The pointwise equality of the spectral density sum (3.8) then allows derivation of the dynamics in Y_0. For all z, we have:

$$\begin{aligned}
&\mathrm{Var}(\epsilon_0) \cdot (1-z)(1-z^{-1})A_0(z)A_0(z^{-1}) \\
&\quad = \mathrm{Var}(\epsilon) \cdot A(z)A(z^{-1}) - \mathrm{Var}(\epsilon_1) \cdot A_1(z)A_1(z^{-1}) \\
&\quad = \mathrm{Var}(\epsilon) \quad \times \\
&\qquad \left\{ \left(\frac{1}{1-0.44z}\right)\left(\frac{1}{1-0.44z^{-1}}\right) - \left[\frac{1-\gamma}{1-0.44}\right]^2 \cdot \left(\frac{1}{1-\gamma z}\right)\left(\frac{1}{1-\gamma z^{-1}}\right) \right\} \\
&\quad = \mathrm{Var}(\epsilon) \cdot \frac{(1-\gamma z)(1-\gamma z^{-1}) - \left[\frac{1-\gamma}{1-0.44}\right]^2 (1-0.44z)(1-0.44z^{-1})}{(1-0.44z)(1-\gamma z)(1-0.44z^{-1})(1-\gamma z^{-1})}.
\end{aligned}$$

Since the numerator of the right hand side expression vanishes at $z = 1$, we can divide both sides by $(1-z)(1-z^{-1})$ to obtain:

$$\begin{aligned}
&\mathrm{Var}(\epsilon_0) \cdot A_0(z)A_0(z^{-1}) = \mathrm{Var}(\epsilon) \times \\
&\qquad \frac{\left((1-\gamma z)(1-\gamma z^{-1}) - \left[\frac{1-\gamma}{1-0.44}\right]^2 (1-0.44z)(1-0.44z^{-1})\right) / (1-z)(1-z^{-1})}{(1-0.44z)(1-\gamma z)(1-0.44z^{-1})(1-\gamma z^{-1})} \\
&\quad = \mathrm{Var}(\epsilon) \cdot \frac{\gamma - \left[\frac{1-\gamma}{1-0.44}\right]^2 \times 0.44}{(1-0.44z)(1-\gamma z)(1-0.44z^{-1})(1-\gamma z^{-1})}.
\end{aligned}$$

But this last expression is simply the covariance function of a second order autoregressive process:

$$Y_0(t) = (0.44 + \gamma)Y_0(t-1) - (0.44 \cdot \gamma)Y_0(t-2) + \epsilon_0(t),$$

where $\mathrm{Var}(\epsilon_0) = \mathrm{Var}(\epsilon) \cdot \left(\gamma - \left[\frac{1-\gamma}{1-0.44}\right]^2 \times 0.44\right)$ is positive since $0.44 < \gamma < 1$.

Recall the contribution to $\mathrm{Var}(\Delta C)$ is $A_1(\beta)^2 \cdot \mathrm{Var}(\epsilon_1)$ for disturbances ϵ_1 that have permanent impact on labor income, and $(1-\beta)^2 A_0(\beta)^2 \cdot \mathrm{Var}(\epsilon_0)$ for disturbances ϵ_0 with only transitory effects. For the example here, these are:

$$(1 - \gamma\beta)^{-2}(1-\gamma)^2(1-0.44)^{-2} \times 636.1,$$

and:

$$(1-\beta)^2(1-0.44\beta)^{-2}(1-\gamma\beta)^{-2}\left[\gamma - \left[\frac{1-\gamma}{1-0.44}\right]^2 \cdot 0.44\right] \times 636.1,$$

respectively. The predicted variance of ΔC is simply the sum of these. For $\gamma = 0.5$, the implied value of $\mathrm{Var}(\Delta C)$ is 1989; $\gamma = 0.75, \mathrm{Var}(\Delta C) = 1916$; $\gamma = 0.8, \mathrm{Var}(\Delta C) = 1882$; $\gamma = 0.9, \mathrm{Var}(\Delta C) = 1727$; $\gamma = 0.95, \mathrm{Var}(\Delta C) = 1493$; $\gamma = 0.99, \mathrm{Var}(\Delta C) = 1017$; $\gamma = 0.995, \mathrm{Var}(\Delta C) = 1083$.

Allowing agents to distinguish permanent and transitory components in labor income, therefore, can potentially smooth the consumption implied by the PIH. Even when Y_1 is restricted to be a first order autoregression in first differences, consumption volatility can fall by as much as one half over that when agents forecast labor income using only its past history. The intuition of the previous section is therefore correct: altering the cleaving of a fixed outer envelope spectral density S affects the PIH prediction for the volatility of consumption. This consumption smoothing occurs without having to change any of the univariate properties of the labor income process.

To complete the argument, we need to show a cleaving exists that reconciles the actual volatility in consumption with the predicted volatility, taking as given the univariate dynamics in labor income. This existence question can be formulated in terms of an infinite-dimensional optimization problem. Take as given (1) a Wold decomposition (A, ϵ) for ΔY, and (2) a real interest rate r implying a value for β. What is the minimum value of $A_1(\beta)^2 \mathrm{Var}(\epsilon_1) + (1-\beta)^2 A_0(\beta)^2 \mathrm{Var}(\epsilon_0)$ such that (3.8) is satisfied? Formally, we have to solve:

$$\inf_{A_1, A_0, \mathrm{Var}(\epsilon_1), \mathrm{Var}(\epsilon_0)} A_1(\beta)^2 \mathrm{Var}(\epsilon_1) + (1-\beta)^2 A_0(\beta)^2 \mathrm{Var}(\epsilon_0)$$

subject to:

a. $\mathrm{Var}(\epsilon)|A(e^{-i\omega})|^2 = \mathrm{Var}(\epsilon_1)|A_1(e^{-i\omega})|^2 + \mathrm{Var}(\epsilon_0)|1-e^{-i\omega}|^2|A_0(e^{-i\omega})|^2$ for all ω,

b. (A_1, ϵ_1) and (A_0, ϵ_0) are Wold representations.

The natural parameter space is infinite-dimensional and equals $\ell^2 \times \ell^2 \times \mathcal{R}_+^2$. Given (1) and (2), consumption displays excess smoothness if this program has value exceeding the sample estimates of $\mathrm{Var}(\Delta C)$.

Instead of solving this problem directly, it is sufficient to display an example satisfying a. and *b.* that achieves a value equal to the sample estimate of the variance of consumption changes. As before, a choice for A_1 immediately determines all the remaining parameters. Equality of S_1 and S at frequency zero fixes the innovation variance $\mathrm{Var}(\epsilon_1)$ in the permanent component. The pointwise equality for all ω, given as $|1-e^{-i\omega}|^2 S_0(\omega) = S(\omega) - S_1(\omega)$ determines the function S_0. This spectral density S_0 can then be factored to obtain uniquely $\mathrm{Var}(\epsilon_0)$, and A_0 in ℓ^2, where $A_0(0) = 1$, and $A_0(z) \neq 0$ for all $|z| < 1$. This calculation is standard since the resulting spectral density is, by construction, a rational function: see, for example, Rozanov [1967] Chapter 1, Section 10, pp. 43-50.

More explicitly, fix a candidate Wold representation (A, ϵ) for ΔY. Consider lag distributions A_1 of the form:

$$A_1(z) = (1+z)^q / A_{1d}(z),$$

where A_{1d} is some fixed polynomial, such that $A_{1d}(0) = 1$, and $A_{1d}(z) \neq 0$ for $|z| \leq 1$.[18] This restricts the permanent component Y_1 to be an *ARIMA* process, with the moving average part having binomial coefficients. As q increases, the spectral density $\mathrm{Var}(\epsilon_1)|A_1(e^{-i\omega})|^2$ is guaranteed eventually to be bounded from above by any fixed spectral density that shares the same value at frequency zero.[19] By $S_1(0) = S(0)$, it follows that:

$$\mathrm{Var}(\epsilon_1) = 4^{-q} A_{1d}(1)^2 A(1)^2 \times \mathrm{Var}(\epsilon).$$

In the second step, the dynamics of ΔY_0 satisfies:

$$\mathrm{Var}(\epsilon_0) \cdot (1-z)(1-z^{-1}) A_0(z) A_0(z^{-1}) = \mathrm{Var}(\epsilon) \times \left(A(z)A(z^{-1}) - 4^{-q} A_{1d}(1)^2 A(1)^2 \cdot \frac{(1+z)^q (1+z^{-1})^q}{A_{1d}(z) A_{1d}(z^{-1})} \right).$$

The right hand side above vanishes at $z = 1$, by our choice of $\mathrm{Var}(\epsilon_1)$. We can therefore write:

$$\mathrm{Var}(\epsilon_0) A_0(z) A_0(z^{-1}) = \mathrm{Var}(\epsilon) \times \frac{\left(A(z)A(z^{-1}) - 4^{-q} A_{1d}(1)^2 A(1)^2 \cdot \frac{(1+z)^q (1+z^{-1})^q}{A_{1d}(z) A_{1d}(z^{-1})} \right)}{(1-z)(1-z^{-1})}.$$

For sufficiently large q, the right hand side is the covariogram of a real covariance stationary process. We can, therefore, factor it to obtain $\mathrm{Var}(\epsilon_1)$ and A_0 such that (1) $A_0(0) = 1$, (2) the power series expansion of $A_0(z)$ is one-sided in non-negative powers of z, and finally, (3) $A_0(z) \neq 0$ for all $|z| < 1$.[20] Finally, taking a value for β, the resulting lag distributions and innovation variances can be used to find the implied variance of consumption changes.

We fix β to the value implied by a risk-free interest rate r of 1 percent per quarter.[21] Tables 1-9 display the results of the above procedure for each of the candidate Wold representations for ΔY in West [1988]. West estimated a variety of models to check the robustness of his findings. Although our results lead to the opposite conclusion from his, here, as in his work, the findings do turn out to be insensitive to the exact parametric specification. The first panel in each table gives point estimates for alternative *ARMA* parametrizations of the Wold representation of ΔY. In our notation, the *ARMA* parameters ϕ and θ satisfy $A(z) = (1 - \phi_1 z - \phi_2 z^2)^{-1}(1 + \theta_1 z + \theta_2 z^2)$. The innovation variance associated with the given *ARMA* parametrization is next presented. These always exceed 600; the largest value is naturally that in Table 1—the random walk case. Next, LR denotes the implied measure of long run persistence, as, for example, in Campbell and Mankiw [1987]. Following that, ψ_0 is the square root of the ratio of PIH-predicted Var(ΔC) to Var(ϵ); finally, ψ is the actual square root of the ratio of Var(ΔC) to Var(ϵ) found in the data. The discrepancy between ψ_0 and ψ is one representation of the Deaton paradox.

The second panel in each table shows the implied variance of consumption changes due to the hypothesized permanent and transitory components in labor income. For alternative settings of q—the moving average length in ΔY_1—we show first the individual variance contributions of the different kinds of disturbances, and then the sum of these contributions. Notice that the contribution of ϵ_1 is always much larger than that of ϵ_0. This is consistent with the message in Lucas [1987, Chapter 3] that cyclical fluctuations, by comparison with secular movements, are simply not significant for many economic questions.

The last row in this panel shows ψ_1 the square root of the ratio of the implied Var(ΔC)

to Var(ϵ). As q increases, the value of ψ_1 falls monotonically. In the last column of this panel, we show the value of q that implies $\psi_1 = \psi$. Finally, the last section of each table presents the value for Var(ϵ_0) associated with that q that matches ψ_1 to ψ.[22]

The last column in the second panel of each table therefore answers positively the existence question above. For all 9 *ARMA* models hypothesized for the Wold representation of ΔY, there exists a permanent-transitory decomposition that exactly matches the PIH predicted consumption volatility with the data. For all 9 models considered, the long-run measure of persistence *LR* is substantial. Despite this, for all 9 models, the PIH—properly considered—does *not* predict consumption volatility exceeding that in the data. Along this dimension, therefore, the PIH is not inconsistent with the data. In summary, the PIH predictions do not particularly depend on (1) whether labor income is best characterized as being integrated, nor (2) the magnitude of labor income's long run persistence, nor (3) the exact form of labor income's univariate dynamics. By the nature of the argument, it should be clear that these conclusions would hold regardless of the exact parametrization of the Wold representation for ΔY, even beyond the *ARMA* (2, 2) cases explicitly considered here.

Are the permanent-transitory decompositions that reconcile the PIH with the data *reasonable*? It is difficult to interpret directly the permanent and transitory components used here for forecasting, since they are not necessarily structural economic disturbances. Thus, one should not read the q values in the last column of the tables to say that economic agents perceive structural shocks with permanent effects as very long *ARIMA* processes. The true structural shocks agents see are likely to be imperfectly correlated across distur-

bances with permanent and transitory effects.

It might be interesting to explicitly identify the structural disturbances that agents see driving labor income. However, such an exercise is not at all relevant in the current context. Instead, here we might simply compare our transitory component in labor income with stationary components that others have estimated. Figure 2 plots the response in labor income to a unit disturbance in the transitory component. The response for each of the 9 models considered in the tables is graphed. In every case, the effects have a hump shape and decay rapidly: no more than half the original impact of the disturbance remains after four years. This is quite consistent with the moving average representation that, for instance, Blanchard and Quah [1989] call the dynamic effects of "aggregate demand" in their study of GNP.

5. Effects of Agents' Superior Information

Campbell and Deaton [1989] and West [1988] have also emphasized that economic agents are likely to have more information than the econometrician. Equilibrium consumption is then necessarily smoother than when agents use only the past history of labor income to forecast future labor income. The question becomes, could this superior information effect suffice to account for the observed smoothness in the data? Campbell and Deaton [1989] and West [1988] find the answer to be no.

Suppose a researcher attempts to account for this superior information by studying the history of time series observations on consumption and income.[23] Recall that consumption is a martingale under the rational expectations version of the PIH. Thus it is natural to suspect that the history of consumption should contain all the relevant information that

agents use to forecast the future. In other words, even though economic agents are likely to have more information than the researcher, studying the joint consumption-income process should allow discovery of the correct relation between "news" and the reaction in consumption, even though the researcher never directly observes "news." This argument appears to be related to a result in Hansen and Sargent [1981]. Hansen and Sargent's [1981] Theorem shows that, under certain conditions, the hallmark rational expectations cross-equation restrictions hold, even when the researcher uses an information set strictly smaller than that of economic agents.

The explanation given in this paper of course says that agents have more information than the econometrician. Why, then, doesn't consumption appropriately reveal the news that agents see? The reason for this is interesting in its own right: the PIH turns out to imply that agents observe innovations that are not fundamental for the joint consumption-income process. We now show this explicitly.[24]

Under the PIH, and our assumptions on permanent and transitory fluctuations, the joint process for the changes in labor income and consumption is:

$$\begin{pmatrix} \Delta Y(t) \\ \Delta C(t) \end{pmatrix} = \begin{pmatrix} A_1(L) & (1-L)A_0(L) \\ A_1(\beta) & (1-\beta)A_0(\beta) \end{pmatrix} \begin{pmatrix} \epsilon_1(t) \\ \epsilon_0(t) \end{pmatrix}. \tag{5.1}$$

The determinant of this matrix moving average is the function:

$$\delta(z) = (1-\beta)A_0(\beta)A_1(z) - A_1(\beta)(1-z)A_0(z).$$

Two features of this function should be noted here: (1) The determinant δ is different from zero at $z = 1$. The spectral density of the jointly covariance stationary vector $(\Delta Y, \Delta C)$ is therefore full rank at frequency zero. In words, labor income and consumption are *not* cointegrated. (Campbell [1987] has made the same observation.)

The intuition is straightforward: the martingale consumption implication of the PIH means that *any* news will have permanent impact on the level of consumption. In particular, even news that has only transitory effects on labor income has permanent effects on consumption. (2) The determinant δ vanishes at $z = \beta$, which is strictly *inside* the unit circle. But then, ϵ_1 and ϵ_0, as well as all linear combinations of them, cannot be recovered from observations on current and lagged values of the exogenous and endogenous variables ΔY and ΔC (see for example Rozanov [1967], Remark 3, p. 63.).

Note this does not mean agents are somehow forecasting using future values of labor income. Recall agents use only current and lagged values of Y_1 and Y_0 to predict future labor income. The nonfundamentalness means simply that the observed variables contain strictly less information than that in the ϵ's. We can see explicitly the implications for inference by considering a simple example.

Consider the Friedman-Muth model (Muth [1960]): Suppose the permanent component is a random walk, $A_1(z) = 1$, and the transitory component is white noise, $A_0(z) = 1$. Further, assume the innovations ϵ_1 and ϵ_0 have unit variances. Substituting into (5.1), these assumptions imply agents in the economy observe the bivariate income-consumption model:

$$\begin{pmatrix} \Delta Y(t) \\ \Delta C(t) \end{pmatrix} = \begin{pmatrix} 1 & (1-L) \\ 1 & (1-\beta) \end{pmatrix} \begin{pmatrix} \epsilon_1(t) \\ \epsilon_0(t) \end{pmatrix}. \tag{5.2}$$

When an econometrician studies the history of observations on labor income and consumption, the most information she can recover is their true Wold representation—the projection of $(\Delta Y, \Delta C)$ on its lagged values.[25] It is not hard to show, given (5.2), that the unique Wold representation for $(\Delta Y, \Delta C)$, with a pairwise orthogonal unit variance

innovation vector is:

$$\begin{pmatrix} \Delta Y(t) \\ \Delta C(t) \end{pmatrix} = \lambda(\beta)^{-1} \begin{pmatrix} (2-\beta)(1-\frac{1-\beta}{2-\beta}L) & (1-\beta L) \\ \lambda(\beta)^2 & 0 \end{pmatrix} \begin{pmatrix} \eta_1(t) \\ \eta_2(t) \end{pmatrix}, \tag{5.3}$$

where $\lambda(\beta)^2 = 1 + (1-\beta)^2$. The disturbances η_1 and η_2 are pairwise orthogonal white noise having unit variance.[26] While the determinant of the matrix moving average in the agents' model (5.2) is $z - \beta$, and vanishes at $\beta < 1$, that in the econometrician's model (5.3) is $\lambda(\beta)(1 - \beta z)$, which vanishes nowhere on the unit disk. The econometrician's representation is therefore fundamental.

The cross equation restrictions in (5.1) completely describe the predicted response of consumption to "news" in labor income. A disturbance to labor income, whose first difference has Wold lag distribution $(1 - z)A_j(z)$ should lead to a consumption response of $(1 - \beta)A_j(\beta)$. Consider the econometrician's representation (5.3). In response to an η_1 disturbance, consumption responds by $\lambda(\beta)$, which is exactly $\lambda(\beta)^{-1}(2 - \beta)(1 - \frac{1-\beta}{2-\beta}z)$ evaluated at $z = \beta$. Thus, the econometrician will infer that consumption appears to respond appropriately to the disturbance η_1. Next, consider an η_2 disturbance. The econometrician reasons that consumption should respond to η_2 by $\lambda(\beta)^{-1}(1 - \beta^2) > 0$; however, in the data, consumption does not at all react to η_2. In other words, consumption will appear to be "excessively smooth," even though the joint income-consumption process satisfies the PIH.

This example shows why an econometrician, studying the past history of observed model variables, might not draw correct inference on the dynamic effects of different disturbances.[27] It is not the case that agents observe *future* events that the econometrician need only wait to similarly observe. Agents in the model condition their actions

only on observations on the *past* history of permanent and transitory disturbances; no future information is involved.

From (5.1), our theory clearly implies restrictions across the equations for income and consumption. In principle, the model (5.1) could be estimated and tested, for example, by maximizing the Whittle frequency domain likelihood. The results of that exercise are known in advance though. Aggregate consumption is not a martingale, and so the PIH would be rejected: see, for example, Christiano, Eichenbaum and Marshall [1987], Heaton [1988], or Nelson [1987]. But evidently this would be for reasons other than "excess smoothness."

6. Conclusion

Deaton [1987] noted that (1) the permanent income hypothesis, (2) a unit root for the labor income process, and (3) the estimated univariate dynamics of US aggregate labor income, together imply volatile consumption. Such volatility is not seen in the time series data on aggregate consumption. However, if labor income is modelled as trend-stationary, then the permanent income hypothesis implies consumption volatility roughly in line with that in the data. Deaton's finding makes a strong case that the univariate dynamic properties of labor income—whether labor income is difference-stationary or trend-stationary, its long run persistence, its univariate short run dynamics—are relevant for evaluating an important economic hypothesis. More generally, it argues for the importance of measures of persistence, as, for instance, articulately proposed in Campbell and Mankiw [1987].

This paper has shown why the reasoning above is misleading. By making the plausible assumption that agents observe different kinds of disturbances to their labor income stream,

the volatility predictions of the permanent income hypothesis can be brought back firmly in line with the data. This can be done *regardless of the precise form of the univariate dynamics in labor income.* Quite generally therefore, this paper argues that the univariate characterizations of aggregate time series are simply not informative for economic theory.

The idea that there are permanent and transitory disturbances in time series is an old one, going back at least to Milton Friedman. This assumption raises interesting testable hypotheses in many areas of empirical time series research. That agents "see" things unobservable to the econometrician has already shed many useful insights, such as the notion of human capital in labor economics and growth. In the current paper, it has served to explain a puzzle, where consumption smoothness seemed to be inconsistent with a unit roots representation for labor income dynamics.

It is important to emphasize what the paper does not do. First, the unit roots hypothesis has been critical in re-orienting econometric inference and modelling. It has provided rich insights for reinterpreting evidence on many interesting economic propositions.[28] This paper does not, at all, argue against this. The results in this paper do, however, lead one to be extremely skeptical of conclusions such as in Deaton [1987] and Campbell and Mankiw [1987]. The focus in those papers on "large" versus "small" unit roots—an unfortunate terminology introduced in Cochrane [1988]—and the idea that, somehow, this has something to do with interesting economic hypotheses appears unjustified. It would be interesting to display an explicit economic model where this faith is, in fact, well-placed.

Second, the paper does not say that the permanent income hypothesis accurately describes the aggregate time series. The martingale predictions for consumption, originally

developed in Hall [1978], are now well-known to be false in the data. Allowing agents to observe permanent and transitory disturbances separately does not alter this conclusion.

References

Beveridge, Stephen and Nelson, Charles R. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'." *Journal of Monetary Economics* 7 (March 1981): 151-174.

Blanchard, Olivier Jean and Quah, Danny. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* 79 (September 1989): 655-673.

Blinder, Alan S. and Deaton, Angus. "The time series consumption function revisited," *Brooking Papers on Economic Activity* 2 (1985): 465-511.

Caballero, Ricardo J. "Consumption Puzzles and Precautionary Savings," Working Paper 399. New York: Columbia University, August 1988 *(a)*.

Caballero, Ricardo J. "Intertemporal Substitution and Consumption Fluctuations," Working Paper 409. New York: Columbia University, October 1988 *(b)*.

Campbell, John Y. "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis." *Econometrica* 55 (November 1987): 1249-1274.

Campbell, John Y. and Deaton, Angus. "Why is consumption so smooth?" *Review of Economic Studies* (1989), in press.

Campbell, John Y. and Mankiw, N. Gregory. "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics* 102 (November 1987): 857-880.

Christiano, Lawrence J. "Why is Consumption Less Volatile than Income?" *Federal Reserve Bank of Minneapolis Quarterly Review* 11 (Fall 1987); 2-20..

Christiano, Lawrence J. and Eichenbaum, Martin. "Unit Roots in Real GNP — Do We Know and Do We Care?" Mimeographed. Carnegie-Rochester Public Policy Conference, April 1989.

Christiano, Lawrence J., Eichenbaum, Martin, and Marshall, David. "The Permanent Income Hypothesis Revisited." Working Paper no. 60. Minneapolis: Federal Reserve Bank of Minneapolis, 1987.

Clarida, Richard H. "Smooth Consumption and Noisy Permanent Income: Is There a Puzzle?" Mimeographed. New York: Columbia University, June 1988.

Cochrane, John. "How Big is the Random Walk in GNP?" *Journal of Political Economy* 96 (October 1988): 893-920.

Deaton, Angus S. "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?" 121-48 in *Advances in Econometrics, Fifth World Congress*, Vol. II, edited by Truman F. Bewley. New York: Cambridge University Press, 1987.

Diebold, Francis X. and Rudebusch, Glenn D. "Is Consumption too Smooth? Long Memory and the Deaton Paradox." Finance and Economics Discussion Series. Washington D.C.: Federal Reserve Board, March 1989.

Flavin, Marjorie. "The adjustment of consumption to changing expectations about future income," *Journal of Political Economy* 89 (October 1981): 974-1009.

Flavin, Marjorie. "The Excess Smoothness of Consumption: Identification and Interpretation." Working Paper No. 2807. Cambridge: National Bureau of Economic Research, December 1988.

Gali, Jordi. "Finite Horizons, Life Cycle Savings and Time Series Evidence on Consumption" Mimeographed. New York: Columbia University, August 1989.

Hall, Robert E. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy* 86 (December 1978): 971-987.

Hansen, Lars Peter. "Calculating Asset Prices in Three Example Economies." 207-243 in *Advances in Econometrics, Fifth World Congress*, Vol. I, edited by Truman F. Bewley. New York: Cambridge University Press, 1987.

Hansen, Lars P., Roberds, William, and Sargent, Thomas J. "Time Series Implications of Present Value Budget Balance and of Martingale Models of Consumption or Taxes." Mimeographed. Chicago: University of Chicago, 1989.

Hansen, Lars Peter and Sargent, Thomas J. "Formulating and Estimating Dynamic Linear Rational Expectations Models." *Journal of Economic Dynamics and Control* 2 (February 1980): 1-46.

Hansen, Lars Peter and Sargent, Thomas J. "Exact Linear Rational Expectations Models: Specification and Estimation." Research Department Staff Report 71. Minneapolis: Federal Reserve Bank of Minneapolis, 1981.

Heaton, John. "The Interaction between Time-Nonseparable Preferences and Time Aggregation." Mimeographed. Chicago: University of Chicago, November 1988.

Lucas, Robert E. *Models of Business Cycles*. New York: Basil Blackwell, 1987.

Muth, John. "Optimal Properties of Exponentially Weighted Forecasts." *Journal of the American Statistical Association* 55 (June 1960): 299-306.

Nelson, Charles. "A Reappraisal of Recent Tests of the Permanent Income Hypothesis." *Journal of Political Economy* 95 (June 1987): 641-646.

Nelson, Charles and Plosser,Charles. "Trends and random walks in macroeconomic time series." *Journal of Monetary Economics* 10 (September 1982): 139-162.

Quah, Danny. "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds," Working Paper 498. Cambridge: Massachusetts Institute of Technology, March 1989.

Rozanov, Yu A. *Stationary Random Processes*, San Francisco: Holden-Day, 1967.

Sargent, Thomas J. "Two Models of Measurements and the Investment Accelerator." *Journal of Political Economy* 97 (April 1989): 251-287.

Sims, Christopher A. "Bayesian Skepticism on Unit Root Econometrics." *Journal of Economic Dynamics and Control* 12 (June 1988): 463-474.

Stock, James H. (1988). "A Reexamination of Friedman's Consumption Puzzle." *Journal of Business and Economic Statistics* 6 (October 1988): 401-407.

Stock, James H. and Watson, Mark. " Variable Trends in Economic Time Series." *Journal of Economic Perspectives* 2 (Summer 1988): 147-174.

Watson, Mark. "Univariate detrending methods with stochastic trends." *Journal of Monetary Economics* 18 (July 1986): 49-75.

West, Kenneth D. "The Insensitivity of Consumption to News about Income." *Journal of Monetary Economics* 21 (January 1988): 17-33.

Wilson, G.T. "Factorization of the Covariance Generating Function of a Pure Moving Average Process." *SIAM Journal of Numerical Analysis* 6 (March 1969): 1-7.

Footnotes

* Department of Economics, MIT and NBER. I am grateful to Ricardo Caballero, John Campbell, Larry Christiano, Angus Deaton, Marjorie Flavin, Kenneth West and David Wilcox for discussions and correspondence. An anonymous referee of this Journal and seminar participants at Columbia, Berkeley, MIT, the NBER Summer Institute, Stanford and UC San Diego also provided helpful comments. I thank the MIT Statistics Center for its hospitality. All errors and misinterpretations are mine.

[1] It is, of course, Nelson and Plosser [1982] who have forcefully drawn macroeconomists' attention to the fact that many aggregate time series may be difference stationary rather than trend stationary. In Nelson and Plosser's terminology, a *difference stationary* series has first difference that is covariance stationary, although the series itself is not; a *trend stationary* series is covariance stationary about a deterministic time trend. In econometric terminology therefore, a difference stationary series is integrated of order 1, or simply integrated, when the order can be omitted without ambiguity.

[2] This characterization is explicitly derived from an optimizing equilibrium model below. Thus, it should be distinguished from that such as in Stock [1988], where consumers ignore transitory income altogether. Further, the coefficient on permanent income in the consumption equation in that work is treated as a free parameter. By contrast, in the kind of models considered here, that coefficient is intimately related to labor income dynamics. Stock [1988] expertly applies recent developments in the theory of regression with cointegrated variables to re-establish Friedman's assertions about errors-in-variables bias. Deaton's paradox does not arise in those kinds of models. The questions of interest there

are different from those considered here.

[3] Recent papers that have provided arguments similar in spirit to that here are Blanchard and Quah [1989] and Christiano and Eichenbaum [1989]. The first does so in a Keynesian model with sticky nominal wages, while the second makes the point in discussing productivity disturbances in a real business cycle model.

[4] Note that we are not suggesting that this permanent-transitory decomposition will explain the other anomalies in the predictions of the permanent income model. For instance, it is now well-known that the martingale implication for consumption is simply false in aggregate time series data. See among others Caballero [1988a, b], Christiano, Eichenbaum and Marshall [1987], Heaton [1988] and Nelson [1987].

[5] After I had completed a first draft of this paper, Lars Hansen, John Heaton, and Thomas Sargent pointed out to me that Hansen, Roberds and Sargent [1989] contains analogous results.

[6] See the quote in the introduction of this paper. Their models however also imply consumption is not a martingale.

[7] We assume throughout that expectations coincide with linear projections.

[8] Loosely speaking, the annihilation operator modifies its operand by removing the part in strictly negative powers of L in the operand's Laurent series expansion. See Hansen and Sargent [1980].

[9] Here we need to calculate the expected present discounted value of a process Y that is not stationary. Notice that the resulting expression contains a singularity on the unit

circle. However, the present discounted value turns out, nevertheless, to be well-defined, by the reasoning surrounding equation (A3) of Appendix A in Hansen and Sargent [1980].

[10] See Deaton [1987], and among others, Campbell and Deaton [1989] and West [1988]. This result is remarkably robust across alternative specifications for A; see Diebold and Rudebusch [1989] and West [1988].

[11] This unit root characterization will be maintained in the subsequent analysis since "excess smoothness" arises only in this case.

[12] Quah [1989] shows how to construct such a decomposition where one component is integrated, the other is stationary, and the innovations in the two components are uncorrelated at all leads and lags. That paper also proves such an orthogonal decomposition can always be found. This is unlike the orthogonal decomposition in Watson [1986] which, under some circumstances, may not exist. It is, however, a maintained assumption that in the economy there are different structural disturbances, not perfectly correlated, that have permanent and transitory effects on labor income. Finally, note that in the current context, the Beveridge-Nelson decomposition is not an interesting one to consider: When the two components are perfectly correlated, forecasts of future labor income are invariant to whether we use the Beveridge-Nelson decomposition or the univariate Wold representation.

[13] From the earlier footnote, this orthogonality assumption is without loss, as the representation is to be used only in forecasting future labor income. Since the structural disturbances to labor income may be correlated, Y_1 and Y_0 may not be directly interpretable. A moment's reflection shows that this does not affect forecasts of future Y, and therefore does not change our predictions for consumption behavior.

[14] Larry Christiano suggested this terminology.

[15] Further, this long run invariance can be shown to hold even when the permanent and transitory components are correlated. See Cochrane [1988] for the case when the permanent component is restricted to be a random walk, and Quah [1989] for the general case.

[16] These numbers are for the Blinder-Deaton [1985] data, which is that typically used in studies on consumption volatility. It is evident that alternative "reasonable" values of r do not fundamentally narrow this difference between predicted and actual variances. Properly accounting for the sampling properties of these estimates also does not alter the conclusion that consumption appears too smooth compared to the predictions of the model; see West [1988].

[17] Technically, a spectral density vanishing at frequency zero does not imply the associated stochastic process is the first difference of another that is covariance stationary. Quah [1989] provides regularity conditions for this implication to hold. It is easy to verify these conditions are satisfied here.

[18] We take the polynomial $A_{1d}(z)$ here to be $(1-0.8z)(1-0.85z)$. This fixes the dominant root in the AR part of A_0 at 0.85. If this isn't done, it might seem like the procedure simply trades off a declining importance in the permanent component for the transitory component approaching non-stationarity. The exact choice, however, is arbitrary otherwise.

[19] This will then satisfy condition a. The moving average form here is also known to minimize the innovation variance of a process that has spectral density fixed at frequency zero; see Quah [1989]. However, this second fact is not directly useful here.

[20] The lag distribution A_0 is, in fact, simply the series expansion of a rational function. The denominator and numerator parts can therefore be obtained separately by a standard algorithm, such as that in Wilson [1969].

[21] This is the value that Christiano [1987] uses. West [1988], on the other hand, uses $r = 0.5$ percent. The results do not much depend on which value exactly is assumed, as long as r is strictly positive.

[22] Although not presented here, I have verified that the zeroes of the autoregressive and moving average parts in the transitory component are strictly outside the unit circle. This property is guaranteed by the algorithm used to factor the spectral density (Wilson [1969]). The condition on the moving average zeroes guarantees that the representations for both ΔY_1 and Y_0 are fundamental, which is necessary for applying the Hansen-Sargent formula.

[23] Campbell and Deaton [1989] and West [1988] use more information than this; therefore the statements below do not apply to their work.

[24] This nonfundamentalness is a property of the Hilbert spaces spanned by the history of the ϵ's and that of the observed sequences ΔY and ΔC. A little reflection therefore shows that it is invariant to whether or not the structural shocks to the economy are correlated, or equivalently, whether or not ϵ's are the structural innovations.

[25] We can, without loss, restrict analysis to the second moment properties of the data.

[26] We obtain this by Rozanov's [1967] discussion in Chapter 1, Section 10, pp. 43-50. It is straightforward to verify the matrix covariograms implied by the right hand sides of (5.2) and (5.3) are identical.

[27] This does not contradict Hansen and Sargent's [1981] results. The PIH restrictions on consumption and income are not of the form in their theorem.

[28] For examples, see the excellent paper by Stock and Watson [1988] and references therein. An opposing view is presented in Sims [1988].

Table 1.

AR: ϕ_1	ϕ_2	MA: θ_1	θ_2	$\mathrm{Var}(\epsilon)$	LR	ψ_0	ψ
–	–	–	–	790.9	1.00	1.00	0.56

q	25	50	75	150	200	99
$\mathrm{Var}(\Delta C)$ due to ϵ_1	511.9	399.4	311.7	148.0	90.1	245.6
$\mathrm{Var}(\Delta C)$ due to ϵ_0	4.4	4.8	5.2	6.4	7.1	5.6
Implied $\mathrm{Var}(\Delta C)$	516.3	404.2	316.8	154.4	97.2	251.1
ψ_1	0.81	0.71	0.63	0.44	0.35	0.56

Matching $\mathrm{Var}(\epsilon_0) = 726.5$.

Table 2.

AR: ϕ_1	ϕ_2	MA: θ_1	θ_2	$\mathrm{Var}(\epsilon)$	LR	ψ_0	ψ
0.44	–	–	–	636.1	1.79	1.77	0.62

q	25	50	75	150	200	201
$\mathrm{Var}(\Delta C)$ due to ϵ_1	1312.9	1024.4	799.3	379.7	231.1	228.9
$\mathrm{Var}(\Delta C)$ due to ϵ_0	10.8	11.8	12.8	15.8	17.8	17.8
Implied $\mathrm{Var}(\Delta C)$	1323.7	1036.2	812.1	395.5	248.9	246.7
ψ_1	1.44	1.28	1.13	0.79	0.63	0.62

Matching $\mathrm{Var}(\epsilon_0) = 591.2$.

Table 3.

AR: ϕ_1	ϕ_2	MA: θ_1	θ_2	$\mathrm{Var}(\epsilon)$	LR	ψ_0	ψ
–	–	0.40	–	659.5	1.40	1.40	0.62

q	25	50	75	150	200	151
$\mathrm{Var}(\Delta C)$ due to ϵ_1	836.7	652.8	509.4	242.0	147.3	239.6
$\mathrm{Var}(\Delta C)$ due to ϵ_0	7.1	7.7	8.4	10.3	11.6	10.3
Implied $\mathrm{Var}(\Delta C)$	843.7	660.5	517.7	252.2	158.8	249.9
ψ_1	1.13	1.00	0.89	0.62	0.49	0.62

Matching $\mathrm{Var}(\epsilon_0) = 610.0$.

Table 4.

AR:	ϕ_1	ϕ_2	MA:	θ_1	θ_2	Var(ϵ)	LR	ψ_0	ψ
	0.52	–		-0.10	–	640.6	1.88	1.66	0.62

q	25	50	75	150	200	212
Var(ΔC) due to ϵ_1	1457.7	1137.4	887.4	421.6	256.6	227.8
Var(ΔC) due to ϵ_0	11.8	12.9	14.0	17.4	19.5	20.0
Implied Var(ΔC)	1469.5	1150.3	901.5	438.9	276.2	247.9
ψ_1	1.51	1.34	1.19	0.83	0.66	0.62

Matching Var(ϵ_0) = 595.7.

Table 5.

AR:	ϕ_1	ϕ_2	MA:	θ_1	θ_2	Var(ϵ)	LR	ψ_0	ψ
	0.43	0.01		–	–	643.6	1.79	1.77	0.61

q	25	50	75	150	200	206
Var(ΔC) due to ϵ_1	1328.4	1036.5	808.7	384.2	233.9	220.3
Var(ΔC) due to ϵ_0	10.9	11.9	13.0	16.0	18.0	18.2
Implied Var(ΔC)	1339.3	1048.4	821.7	400.1	251.8	238.5
ψ_1	1.44	1.28	1.13	0.79	0.63	0.61

Matching Var(ϵ_0) = 598.4.

Table 6.

AR:	ϕ_1	ϕ_2	MA:	θ_1	θ_2	Var(ϵ)	LR	ψ_0	ψ
	–	–		0.45	0.11	633.3	1.56	1.55	0.63

q	25	50	75	150	200	169
Var(ΔC) due to ϵ_1	997.6	778.4	607.3	288.5	175.6	238.9
Var(ΔC) due to ϵ_0	8.4	9.2	9.9	12.2	13.7	12.8
Implied Var(ΔC)	1006.0	787.5	617.2	300.7	189.3	251.7
ψ_1	1.26	1.12	0.99	0.69	0.55	0.63

Matching Var(ϵ_0) = 586.9.

Table 7.

AR:	ϕ_1	ϕ_2	MA:	θ_1	θ_2	Var(ϵ)	LR	ψ_0	ψ
	0.86	-0.17		-0.44	–	646.2	1.81	1.79	0.61

q	25	50	75	150	200	209
Var(ΔC) due to ϵ_1	1364.9	1065.0	830.9	394.7	240.3	219.8
Var(ΔC) due to ϵ_0	11.1	12.2	13.2	16.3	18.3	18.7
Implied Var(ΔC)	1376.0	1077.1	844.2	411.0	258.6	238.5
ψ_1	1.46	1.29	1.14	0.80	0.63	0.61

Matching Var(ϵ_0) = 600.9.

Table 8.

AR:	ϕ_1	ϕ_2	MA:	θ_1	θ_2	Var(ϵ)	LR	ψ_0	ψ
	0.65	–		-0.22	-0.12	640.6	1.89	1.86	0.62

q	25	50	75	150	200	213
Var(ΔC) due to ϵ_1	1474.4	1150.4	897.6	426.4	259.6	228.2
Var(ΔC) due to ϵ_0	11.6	12.7	13.9	17.2	19.4	20.0
Implied Var(ΔC)	1486.0	1163.2	911.5	443.6	279.0	248.1
ψ_1	1.52	1.35	1.19	0.83	0.66	0.62

Matching Var(ϵ_0) = 595.5.

Table 9.

AR:	ϕ_1	ϕ_2	MA:	θ_1	θ_2	Var(ϵ)	LR	ψ_0	ψ
	0.50	0.07		-0.07	-0.13	649.2	1.86	1.84	0.61

q	25	50	75	150	200	216
Var(ΔC) due to ϵ_1	1454.5	1134.8	885.5	420.6	256.1	218.5
Var(ΔC) due to ϵ_0	11.6	12.7	13.8	17.1	19.3	20.0
Implied Var(ΔC)	1466.0	1147.6	899.3	437.7	275.3	238.4
ψ_1	1.50	1.33	1.18	0.82	0.65	0.61

Matching Var(ϵ_0) = 603.8.

FIGURE 1

A Spectral Density Cleaving

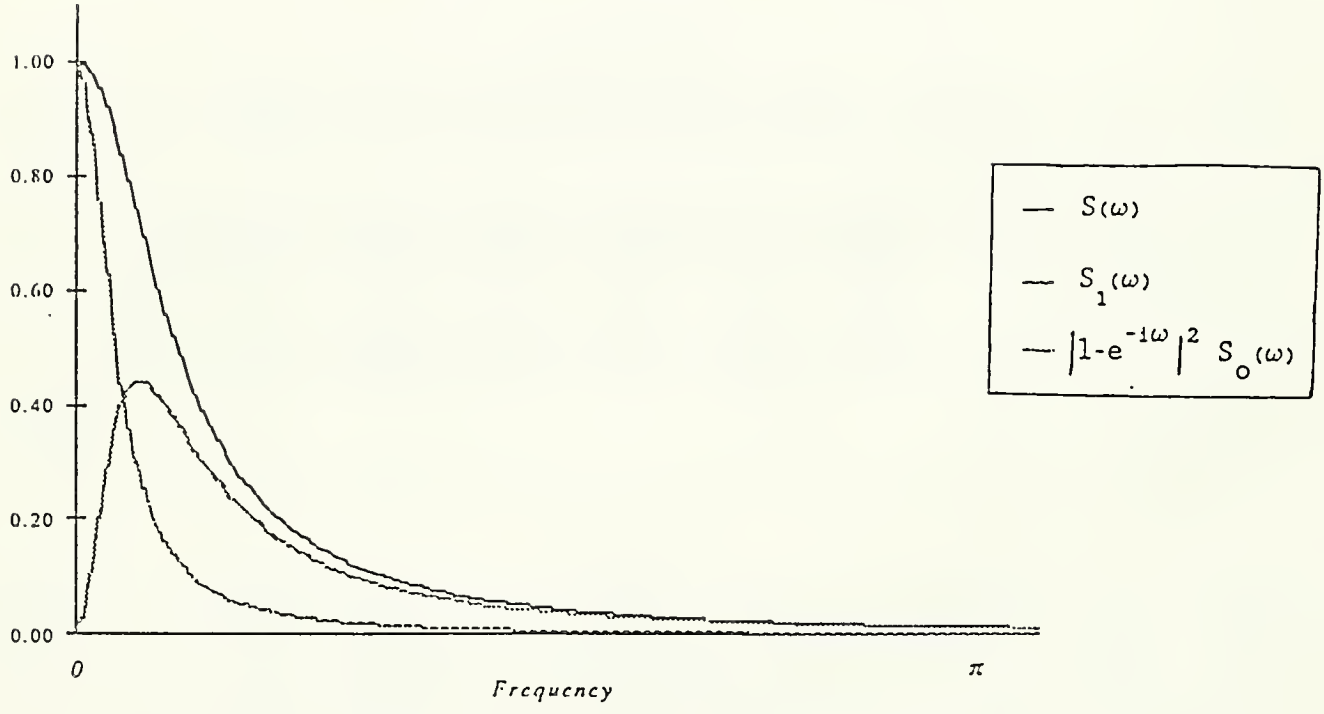
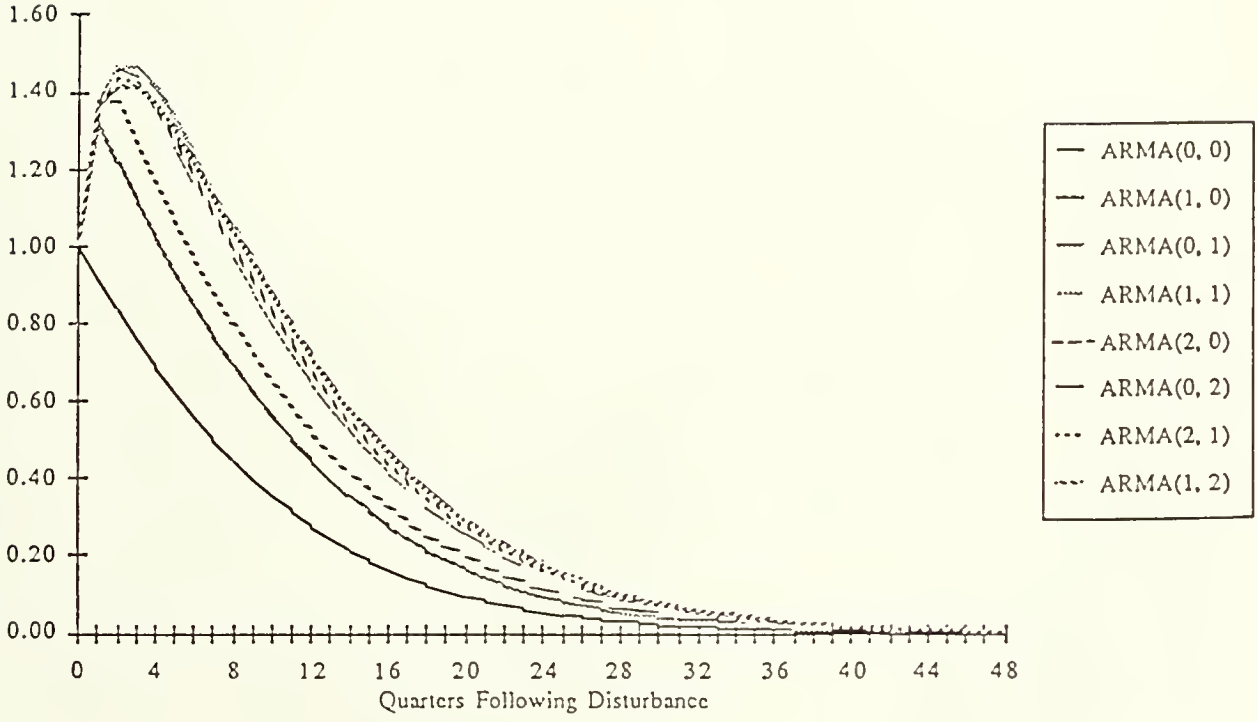


FIGURE 2

Labor Income Response to an Innovation in the Stationary Component



2385 U08

Date Due 2-12-90

Lib-26-67

MIT LIBRARIES DUPL 2

3 9080 00578985 1